UK Parliament
POST

POSTnote 737

By Hannah Gardiner,
Natasha Mutebi
31 January 2025

AI and mental healthcare: opportunities and delivery considerations

Overview

- Mental healthcare provision in the UK faces rising demand and workforce capacity issues. Evidence indicates artificial intelligence (AI) could support administrative tasks and provide direct support to service users.
- Innovations within 'precision psychiatry' appear to show promise for precise diagnosis, monitoring, and risk prediction of mental health conditions.
- Trials and use of AI are ongoing across the UK and internationally, with important distinctions between: AI tools built for mental healthcare in the NHS that are subject to standards and regulation; consumer products for wellbeing with less regulatory oversight; and AI tools not intended for mental health but used by people with mental health challenges.
- Stakeholders suggest delivery requires effective integration into existing care pathways and delivery infrastructure, upskilling staff, enhanced data infrastructure, and appropriate regulation and oversight. Evidence highlights the role of public trust, engagement and education.
- Stakeholders note that high-quality data is needed for product development and to generate robust evidence. They also note the need for longer-term evaluation of safety, efficacy, cost-effectiveness and operational efficiency.
- Ethical, health and legal concerns are also important, see PN738 for information on these challenges and potential regulatory responses.

Background

Reported mental ill-health challenges are increasing[a], with data from the National Health Service (NHS) and the Office for National Statistics (ONS) showing mental ill-health affects approximately 1 in 5 young people[2] and adults.[3] (Figure.1).[4,5] For the month of November 2024, 1.97 million people were in contact with mental health services, with 450,180 new referrals in that month alone.[6]

Mental health services are mostly commissioned at a local level, and funding is not ringfenced, although there is a mental health investment standard as guidance (CBP7547).

Mental healthcare is delivered across a range of settings, including primary care and community services, although the delivery structure differs between areas due to the predominantly local commissioning (CPB7547).

There are notable regional and demographic disparities in who is accessing services[b].

Emerging technology including AI is a UK-wide policy area whereas healthcare is a devolved policy area. This POSTnote refers to mental healthcare delivery in England, although the research described is likely to have broad applicability across the UK.

[a] Discussion around the increase in reported mental health challenges does not draw definitive conclusions on the causation. Possible explanations include: the prevalence of mental ill-health genuinely increasing, increased reporting due to a decrease in stigma, or changes in diagnosis practice having an impact.[1] There are also notable differences in population subgroups such as ethnic groups and gender.[1]

[b] For example, there are regional disparities in waiting times for accessing services and between the first and second treatment, there are also disparities in recovery and improvement rates for some minoritised groups, see detail in CP06988.

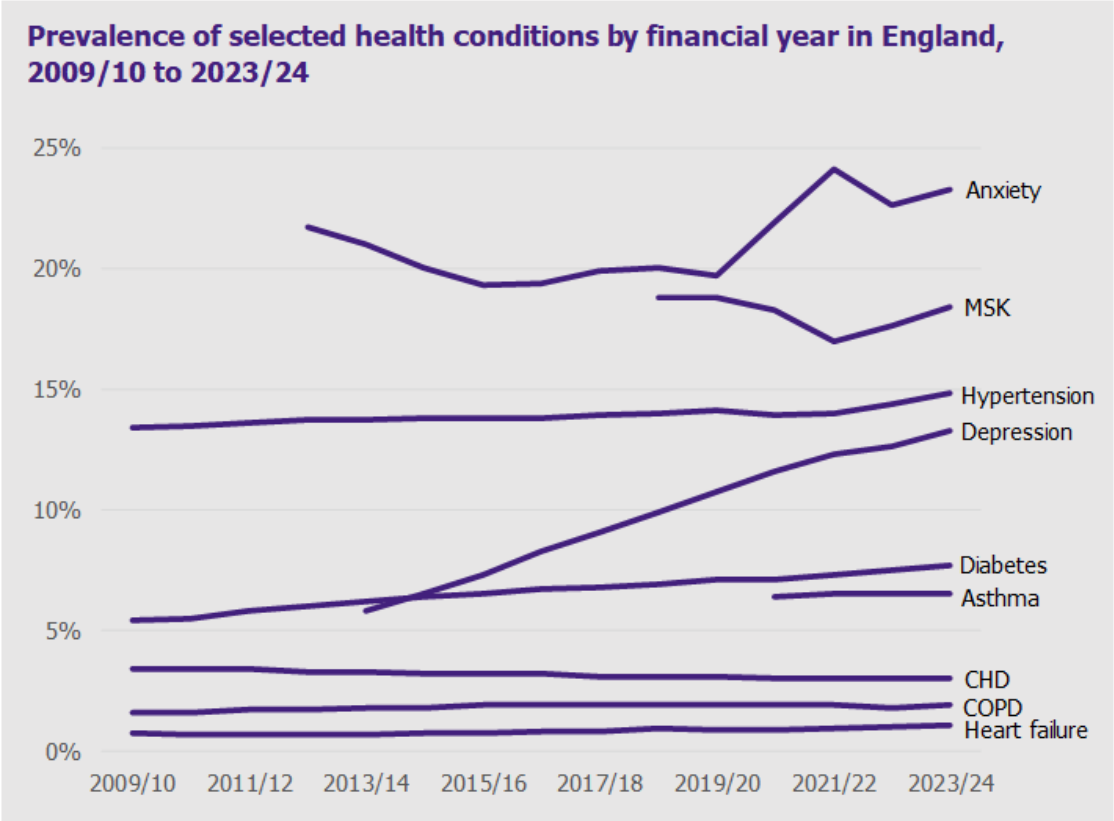

Figure 1: This figure illustrates increasing prevalence of mental health conditions as compared to other long-term conditions. Notes: Depression is reported for those aged 18+, anxiety and MSK are reported for people aged 16+, diabetes for people aged 17+, and asthma for those aged 6+. All other prevalence rates are for all ages. Data for anxiety and MSK are self-reported, rather than medically diagnosed. MSK is reported by calendar year. Abbreviations: Coronary heart disease (CHD), chronic obstructive pulmonary disease (COPD), musculoskeletal disease (MSK). Data Source: DHSC, Fingertips, Health trends in England. Adapted From the Darzi Report (2024, p.23)[5] produced by Sonja Stiebahl (House of Commons Library).

The 2024 independent government commissioned Darzi Report, and other recent reports, outline NHS capacity challenges in the face of rising demand, with struggles to recruit and retain staff. For example, mental health nurse numbers returned to 2009-10 levels[c] in 2023-24.[5,7,8]

The NHS has various strategies to improve capacity issues,[9–11] including digitalisation,[9–12] which is suggested to decrease clinician time spent on administration,[13] or to improve access through digitally-enabled therapy.[13]

A 2020 National Audit Office report on Digital Transformation in the NHS identified constraints including outdated IT systems, lack of skills, and a need for financial investment.[14]

However, a 2022 government consultation stated that digitalisation in healthcare accelerated since the covid-19 pandemic, thus presenting an opportunity to use advanced technologies to support mental healthcare delivery.[15]

[c] The period of 2009-10 was prior to the UK government's austerity programme in which they reduced public spending. See also p.7 of the Darzi report (2024).

Artificial intelligence (AI) tools and services are part of the broader digitalisation of healthcare (PN637).[16] Key terms related to AI are in table 1. Policy activities related to mental health and AI are in table 2.

At present there are several standards, regulations, and guidance that apply to Digital Mental Health Interventions (DMHIs), particularly if they are supplied to the NHS.

Notably, any devices classified as medical devices are subject to additional regulatory oversight. See details of these standards, regulations and guidance in the 'Relevant Policy and Regulation' section in PN738. Work is currently underway to identify whether and how these should be updated given technological advances - see 'Regulatory Challenges and Responses' section in PN738, which also has international comparison case studies.

For detailed definitions of AI technical terms, please consult the POST AI Glossary.

Table 1 Key terms

Term	Description
Artificial intelligence (AI)	AI technologies are tools and services which have some level of autonomy in undertaking activities, generating new predictions and decision-making without direct human control. Some of which can continue to adapt after being 'trained' on datasets. AI is an overarching category, and there are many types of AI technologies, some of which overlap or build on each other. Further detail on AI technologies: 'AI an explainer' PB057
Machine Learning (ML) or Predictive AI	These systems learn to find patterns in training datasets which are then typically applied to new data to make predictions, carry out processing tasks, or provide useful outputs (e.g. text translation or data modelling).
Generative AI (GenAI)	An AI model which generates text, images, audio, video or other media in response to user prompts[d]. These are advanced ML Models (currently the most common are Large Language Models) are trained on large amounts of data and fine-tuned by developers to create new data and outputs with similar characteristics to the data it was trained on. Generative AI can adapt and learn over time.
Rule-based AI	An alternative to GenAI is 'rule-based'[e] AI. Here the system uses a set of predetermined rules to make decisions based on logical reasoning. These are often used in systematic processes or diagnostic settings.[18]

[d] "A prompt is a natural language request submitted to a language model to receive a response back. Prompts can contain questions, instructions, contextual information, few-shot examples, and partial input for the model to complete or continue. From a user prompt, the model responds with generated text, embeddings, code, images, videos, music, and more."[17]

[e] Rule-Based AI systems use a set of predetermined rules to make decisions (e.g. clinical standards), rather than learning from data to make decisions, this can make them more predictable and transparent but makes them less adaptable. See more here.

Opportunities and applications

Recent academic and World Health Organisation (WHO) research highlights multiple ways AI and machine learning (ML) techniques could support mental healthcare delivery,[19–21] including:

- improved diagnosis and triage[f] [23–25]
- improving prediction of individual risks of being affected by conditions such as mood disorders, or developing outcomes such as psychosis or suicidal behaviours[25–28]
- predicting service user response to interventions, such as responding positively to particular antidepressants or psychotherapeutic techniques[25,26,29,30]
- personalising treatment, for example tailoring therapeutic interventions[25,29,31]
- supporting treatment adherence, for example engaging service users between treatments[24,30]
- supporting stratified care[g] approaches[33,34]
- responsive self-management support available on-demand and outside of working hours and matching people with therapists[35,36]
- training or operational support for mental healthcare staff[21,24,25,37–39]

Stakeholders also suggest potential benefits of accessibility, scalability, and consistency in service delivery.[19]

Some stakeholders suggest AI-based tools could support healthcare professionals to provide personalised care more quickly than currently, which could prevent development of severe mental illness outcomes (such as symptoms worsening or relapses), thus alleviating pressure on services.[40] See figure 2.

Other emerging uses of AI include supporting grieving processes,[41–44] aiding social prescribing[h] [46,47] and supporting safeguarding and user engagement within digital peer support platforms.[48,49]

There is investment in DMHI's from both private and public sectors. In 2024 the Wellcome Trust announced funding for large-scale projects (£3m-£7m each) to

[f] Triage is the process of deciding the appropriate next steps for a patient, who is best placed to deliver them, and when they should happen.[22] The NHS has specific guidance for Digitally Enabled Triage.

[g] In stratified care approaches people are stratified by their risk profile and then put on differentiated care pathways, some evidence suggests this could be useful for depression.[32,33]

[h] Social prescribing "connects people to activities, groups, and services in their community to meet the practical, social and emotional needs that affect their health and wellbeing."[45] For example, a person being connected to a local gardening group.

develop scalable DMHIs.[50] Wellcome Trust will also be funding research and development of GenerativeAI models for applications in mental health in 2025.[51,52] See section on 'The Development of GenerativeAI'.

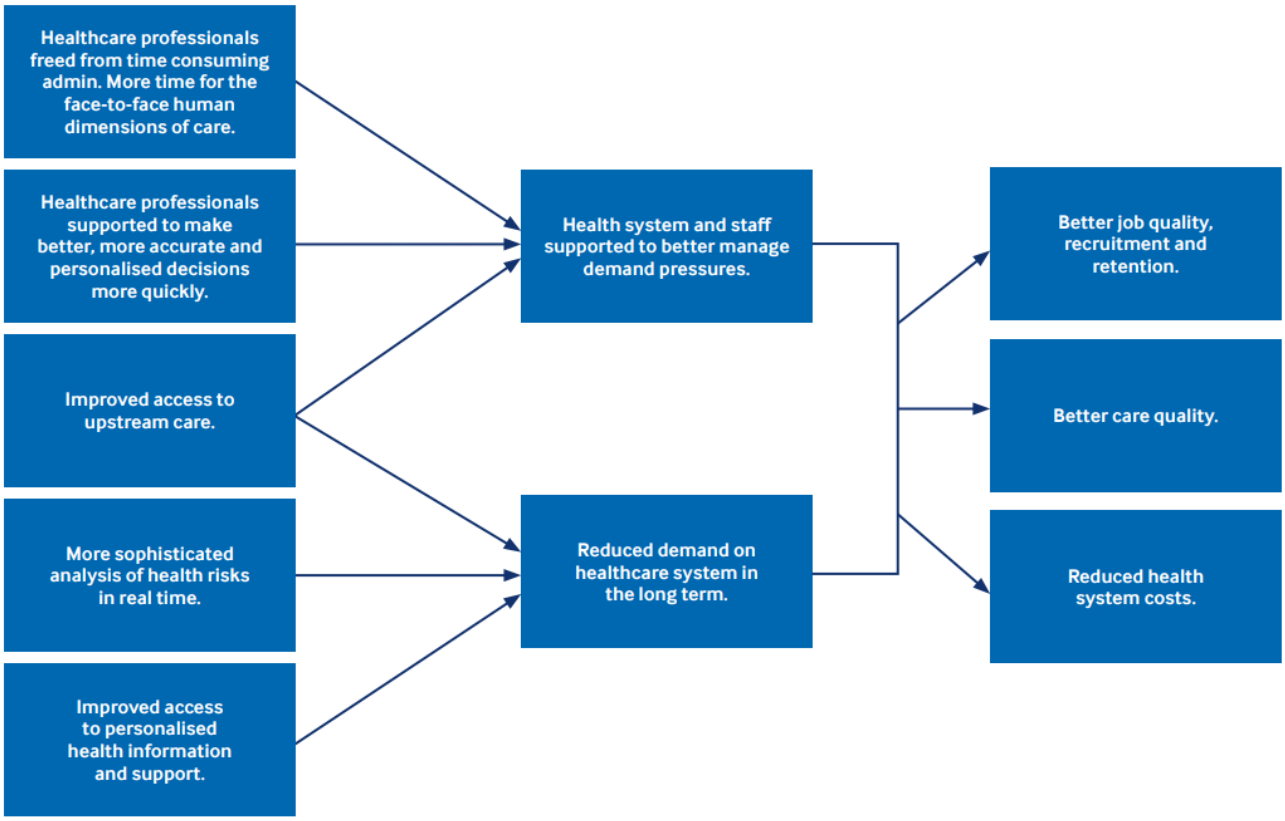

Figure 2: Potential healthcare AI outcomes, extracted from the British Medical Association report 'Principles for Artificial Intelligence (AI) and its application in healthcare' (2024).[38] The NHS make similar suggestions regarding potential benefits.[53]

How autonomous?

Some public engagement and research recommends DHMIs should supplement rather than replace face-to-face delivery,[35,36,54–64] with lived experience perspectives describing human contact as highly valuable.[63]

Research also suggests fully-automated AI tools and services could exist in the future,[65,66] although it highlights that GenAI is unlikely to have sufficient capacity or trust to replace healthcare professionals in the short term.[67] For example, small scale trials indicate fully-automated chatbot DMHI's could reduce symptoms of anxiety and depression.[68–71] Academic analysis suggests such interventions could support people whilst they are on waiting lists.[72] There are already numerous commercial 'wellbeing apps' operating outside the NHS, some of which are fully-automated, although research raises concerns around efficacy, safety and privacy.[73–75]

Research also highlights that work is needed on formation of "digital therapeutic alliances"[i],[65] and to mitigate risks by ensuring human oversight.[66] Academic and lived experience analyses emphasise ensuring crisis support pathways are available for service users of all types of DMHI's.[54,55,76] Youth perspectives on fully-automated chatbot DMHIs emphasise ensuring transparency of whether content is evidence-based, alongside effective privacy, data protection, and consent processes.[76]

There are also legal implications if fully automated decision-making is implemented. The Data Protection Act 2018 specifies when this can be carried out, and what protections must be in place.[77]

Effectiveness for service users

Purpose-built chatbots

Individual trials evaluating tools and some systematic reviews indicate AI chatbots[j] that are purpose-built for mental health intervention could support treatment of specific conditions such as eating disorders or depression. [57,69,71,72,79–86] They could also enable preventative health and wellbeing measures and behaviour change.[70,80,87] For example, an AI-based tool supporting service users between therapy sessions increased their improvement and recovery rates.[88] Individual studies evaluating tools identified convenience, flexibility of access and perceived non-judgemental characteristics as valued by users,[89–91] although engagement relies on self-initiative.[90]

Several meta-analyses[k] conducted in 2023 and 2024 similarly found statistically significant reductions of immediate mental health symptoms, such as anxiety, when people affected by mental health challenges used AI-based chatbot tools. [81,85,86]

However, statistically significant results were not found for overall improvements in psychological wellbeing,[81] nor for long-term effects on most mental health outcomes.[85,86] A 2024 review found mixed results with some studies showing AI chatbot's increased service user engagement while others concluded engagement was challenging.[80]

[i] "digital therapeutic alliance" (DTA), aims to conceptually capture and measure the therapeutic quality of online psychological therapy or digital mental health interventions.

[j] AI chatbots are increasingly improving capabilities in mimicking real conversation and portraying emotions.[78] Other names include smart bots, artificial conversation entities/agents, interactive agents, and digital assistants.[78]

[k] A meta-analysis (plural: meta-analyses) is an academic research study which examines outcomes from a number of studies meeting predefined quality and relevance criteria, and uses a statistical analysis to identify if the outcomes across all of them are statistically and clinically significant. For example, in the case of AI-based chatbot apps, there might be 20 individual studies or trials which each examine the effectiveness of similar tools. The meta-analysis would look across the results of all 20 studies.

Non-purpose-built chatbots

Alongside purpose-built mental health tools aiming to supply the NHS, there are many commercially available AI products that were not designed for mental health purposes and are subject to less regulatory and governance oversight.

Studies suggest commercially available companion chatbots[l] can be useful for early or temporary intervention or wellbeing support,[94–98] particularly for highly lonely populations.[97,99]

However, some evidence shows chatbots could create emotional-dependence, over dependence, or addictive behaviours. They could also cause user distress if updates cause chatbots to react differently.[96,98,100] Some stakeholders argue that creating attachments with chatbots could be detrimental to people forming human relationships.[96,101]

Others report chatbots have helped stimulate rather than displace people forming human relationships,[99] or mixed results[94]. For example, research found chatbots can provide a safe space for people to experiment and build social skills and confidence.[96,98]

Concerns around commercial chatbot use include data protection and privacy and the potential for serious harms to the public including instances of suicide (PN738). [102,103]

AI-driven digital tools

Beyond chatbots, other AI-driven apps interfaces showing promise in early trials include apps supporting users to learn how to reframe negative thoughts[m],[104] and to develop communication and emotional management skills.[70]

See examples of currently available digital tools in Table 3, which include chatbots and other interventions. Please note, although specific services are highlighted as examples, many companies offer more than one service, and many of these tools do not use GenAI but other AI, such as rules-based AI.

Additionally, AI tools built to be supplied to the NHS are subject to various criteria, standards and oversight, particularly if they classified as medical devices (PN738).

[l] AI Companion Chatbots use GenAI, they allow users to customise their own companion whom they can interact with, including use as virtual lovers/partners. They are not specifically designed to be mental health interventions. Specific ethical[92] and privacy[93] concerns have been raised related to these products.

[m] Cognitive reframing supports people to shift their perspectives or how they view situations, people, or relationships and is sometimes used by therapists to support people caught in negative thought patterns.

Evidence evaluation

National Institute for Health and Care Excellence (NICE)[n] Early Value Assessments[o] were completed for DMHIs that aim to support treatment of anxiety,[107] depression,[108] and psychosis.[109]

Most DMHIs with AI/ML integrations were considered only suitable for use within ethically approved research,[110,111] apart from the platform CareLoop[p]. NICE considered CareLoop to have more robust evidence at the time of assessment and it can therefore be used as an option in the NHS whilst further evidence is generated.[112]

Researchers suggest studies over longer periods of time are needed, alongside research on whether and how individual and contextual differences can impact peoples' responses to chatbots,[65,81,86,90,94] and comparisons with DMHIs which do not integrate AI.[85]

Some authors question whether therapeutic relationships can be effective without human connection.[113,114]

Many studies are currently led by or involve commercial entities, see discussion in 'The Quality of Research and Evidence' section in PN738.

Table 2 Examples of purpose-built tools for direct service user support

Provider	What do they offer	Examples of trial/use by UK healthcare providers
Wysa	Provide AI chatbots for several uses including: employee wellbeing, a 'copilot' to support therapists, a referral/triage app, and an app to support people's ongoing engagement in their care pathway.[72,91,115–117]	Lancashire and South Cumbria NHS Foundation Trust, Coventry and Warwickshire Partnership Trust, Central and North West London NHS Foundation Trust, Black Country Healthcare Foundation Trust, Dorset Healthcare NHS Trust, Whittington Health NHS Trust, Alliance Healthcare (Tees Esk and Wear Valleys NHS Trust), Mind (Hammersmith, Fulham, Ealing, Hounslow), Norfolk and Suffolk NHS Foundation Trust, Vita Healthcare

[n] The National Institute for Health and Care Excellence (NICE) carry out assessments of medical treatments, including technology appraisals.[105]

[o] National Institute for Health and Care Excellence (NICE) Early Value Assessments (EVAs) are focused on technologies in priority areas where the NHS needs to address issues such as lengthy waiting lists, such as mental health. EVAs offer guidance to NHS about use of and potential value of technologies, so they can start to benefit from them whilst more evidence is generated.[106]

[p] The CareLoop platform provides healthcare workers real-time symptom insights to inform decision making on treatment and management options. Prediction using proprietary AI within the platform identifies potential markers of escalation for early intervention and prevention.

Iona Mind	Provide low intensity self-paced CBT[q] through an online or mobile phone-based app, with a care team available 24/7. [118]	Unum Group, Aptar Pharma, Office for Veterans' Affairs, NHS Talking Therapies[119]
CareLoop	Provide digital tools including a mobile phone app to predict, avoid and minimise mental health crises.[120–123]	Greater Manchester Mental Health Foundation Trust
IESO	Therapists deliver treatment, supported by online questionnaires to monitor progress, and a chatbot referral form (in partnership with Wysa).[124,125]	NHS Kent and Medway Integrated Care Board, Dorset Healthcare University NHS Foundation Trust, South West London and St Georges Mental Health NHS Trust
Kooth	Provides online peer support through forums, evidence-based resources, and live counselling with human therapists. Has developed an algorithmic tool to supporting content moderation tasks.[49] Anonymised data is also provided to partners to effectively address mental health needs.[49,126–128]	Partnering with NHS trusts, Integrated Care Systems (ICS), and local authorities across England Scotland and Wales (further detail not disclosed)

AI support in health systems operations and training

Researchers, trade unions and the WHO suggest AI technologies could support administrative tasks and medical or nursing education (PN637),[129–132] with unique opportunities for mental health staff training.[24,61,133] Tools are being trialled by many NHS trusts (examples in Table 4, note that there is cross over between these tools and the clinical decision support tools in table 5, and many companies offer more than one service).

NHS Talking Therapies[r] guidance promotes digitalisation but states paper, email and phone services should be maintained.[135] NHS services deploying tools to streamline administration tasks and alleviate staff shortages[132,136] describe adhering to this guidance by maintaining multiple access options.[137]

Industry-led evaluation of implementing an AI referral tool found time savings of 15 minutes per referral (3000 hours pro-rata), and ~20% signposted to more appropriate services, thereby saving additional time.[137] It also identified a 9%

[q] Cognitive Behavioural Therapy (CBT) is a talking therapy which supports people to change the way they think and behave.

[r] The NHS Talking Therapies programme (for anxiety and depression) was formerly known as Improving Access to Psychological Therapies (IAPT). It was "developed to improve the delivery of, and access to, evidence-based, NICE recommended, psychological therapies for depression and anxiety disorders within the NHS"[134]

increase in referrals over a 3 month period (~14,000 service users),[89] and additional benefits of accessibility outside working hours[31] and improved reliable recovery[s] rates of service users.[139]

There is also cross-sector interest in emerging AI-enabled voice recognition technology,[132,140] with hopes to mitigate high workloads and clinician burnout, for example, by products automatically producing draft documentation such as reports.[140]

However, a 2024 review on use of AI across healthcare concluded operation-efficiency increases are inconsistent and more research on this is needed.[141] Healthcare trade unions highlight AI could reduce job pressure, but conversely could create new pressures by reducing 'easy' routine work leaving only complex tasks.[38,132] See also the section on 'AI tools supporting clinical decision making'.

Table 3 Examples of tools to automate and support workflows

Provider name	Provider solution	Examples of trial/use by UK healthcare providers
Censeo	A chatbot to support referral processes. They have also demonstrated feasibility of combining blood biomarkers and questionnaire responses for diagnosis.[142–144]	NHS Hertfordshire Partnership University, Devon Partnership NHS Trust, Sussex Partnership NHS Foundation Trust, Living Well Consortium UK, Mental Health Matters.
Anathem	Their tool uses ambient voice recognition to autogenerate reports from appointments. Preliminary results appear positive but have yet to be published.	Central and North West London NHS Foundation Trust, Great Ormond Street Hospital for Children NHS Foundation Trust and Berkshire Healthcare NHS Foundation Trust.

Precision Psychiatry

Precision psychiatry[t] is an emerging approach that aims to innovate current psychiatry practices by leveraging advances in AI/ML data processing capabilities along with the availability of large and diverse datasets.[146] Academic analysis also suggests AI could contribute to preventative approaches to mental healthcare delivery ('salutogenesis[u])', rather than only treating illness. [147,148,149]

[s] Reliable Recovery is an outcome measure used by the NHS, it means a patient has "moved from being a clinical case at the start of treatment to not being a clinical case at the end of treatment, and there has also been a significant improvement in their condition.[138]

[t] The umbrella term Precision Psychiatry refers to approaches related to four P's of making diagnosis, treatment and prevention "more personalized, proactive, predictive, and precise".[145] See figure 3.

[u] Salutogenesis is a health system theory and orientation to healthcare delivery which argues "that life experiences and how an individual views their life (either positively or negatively) influences their health," takes into account social determinants of health, and encourages health professionals to support service users to strengthen their resources and be empowered to maintain or positively increase their health.[147]

Enhanced detection and monitoring of mental health conditions

Technology enabled data collection

Research highlights smartphones, social media, online platforms and new wearable technology offer possibilities for continuous and passive data collection, enabling service user monitoring for prediction and personalisation of mental health interventions.[35,65,150–166] Digital phenotyping[v] can be delivered by collecting and analysing multiple types of data to monitor people's behaviour.[151,160] At present, commercial products delivering monitoring of mental wellbeing exist,[168–170] though academic studies have raised questions about their accuracy, validity and safety, including privacy concerns.[171–173]

Academic and third sector perspectives highlight opportunities to analyse financial transaction data and identify people who may be struggling with their mental health so proactive interventions can be offered.[174–176] There is a two-way relationship between financial and mental health struggles, and people facing either or both can struggle to ask for support (PN732), so this could reduce barriers to access.

Third sector research into lived experience perspectives found high levels of scepticism about digital phenotyping amongst some, including concerns of threats to data and privacy, and that the approach could lead to service users having less power over their treatment and support.[63] Some academic analysis raises similar concerns.[177]

Using AI to analyse text data

In recent years the capabilities of AI/ML to process text have improved[w], and academic studies highlight new opportunities to analyse text-based data that could enhance delivery of psychiatry practice and carry out research reviews.[178–182] Academic research highlights particular promise from analysing electronic health records[181] or text data produced within controlled environments like online forums or chatbot interactions (as opposed to social media).[180] One UK-based company has already successfully used AI techniques to support moderation of their online peer support forums.[49]

Using diverse data sources to enhance psychiatric practice

AI/ML data processing capabilities can also support diagnosis and prediction (for example, risk of suicide) by analysing large datasets from both clinical testing and real world monitoring using diverse types of data (see Figure 3).[37,153,159,183–189] Academics aim to develop clinically relevant biological or digital

[v] Digital phenotyping is defined as "using data from sensors and interactions on personal digital devices to measure behaviour"[167]

[w] This is related to the development of 'Large Language Models' (LLMs), which are AI models trained on vast amounts of data (Foundation Models) and which leverage advancements in 'Natural Language Processing' (NLP) so they can receive inputs and create outputs in natural written or spoken language. Advancements in NLP have also contributed to opportunities.

biomarkers[x],[151,184,191,192] and redefine the classification, diagnosis and treatment of mental ill heath using these biomarkers.[193]

However, researchers raise concerns about the sample sizes, quality, generalisability, and lack of external validation in current trials of the proposed precision psychiatry methods.[26,146,188,194–197]

Research into lived experience perspectives found some positive outcomes on ML supporting people to understand their own conditions better, or population level insights, but also uncovered concerns around the ability of ML to capture the complexity of human experiences, particularly the wider determinants of (mental) health[y].[63] Another study found service user participants raised concerns about being judged on their genetics (for example their parents' disorders).[199]

In 2022, Wellcome Trust invested £12.5m in the CONNECT study, testing use of ML and active and passively collected data (via wearables and smartphones) to detect early warning signs of relapse into psychosis.[200,201]

In 2024, UKRI funded a project using AI to analyse multiple types of data and novel data collection to support diagnosis and treatment of people with Borderline Personality Disorder.[202,203] A large-scale study exploring detection and treatment of psychosis using AI-tools is also in progress in Germany.[204]

[x] Biomarkers are a measurable indicator of a condition, historically this would have been a biological sample such as a chemical marker in blood. As digitisation has spread across healthcare and health research, the concept of digital biomarkers arose, these are defined "to be a characteristic or set of characteristics, collected from digital health technologies, that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions."[190]

[y] "The wider determinants of health are a diverse range of social, economic and environmental factors which influence people's mental and physical health. Systematic variation in these factors constitutes social inequality, an important driver of the health inequalities"[198]

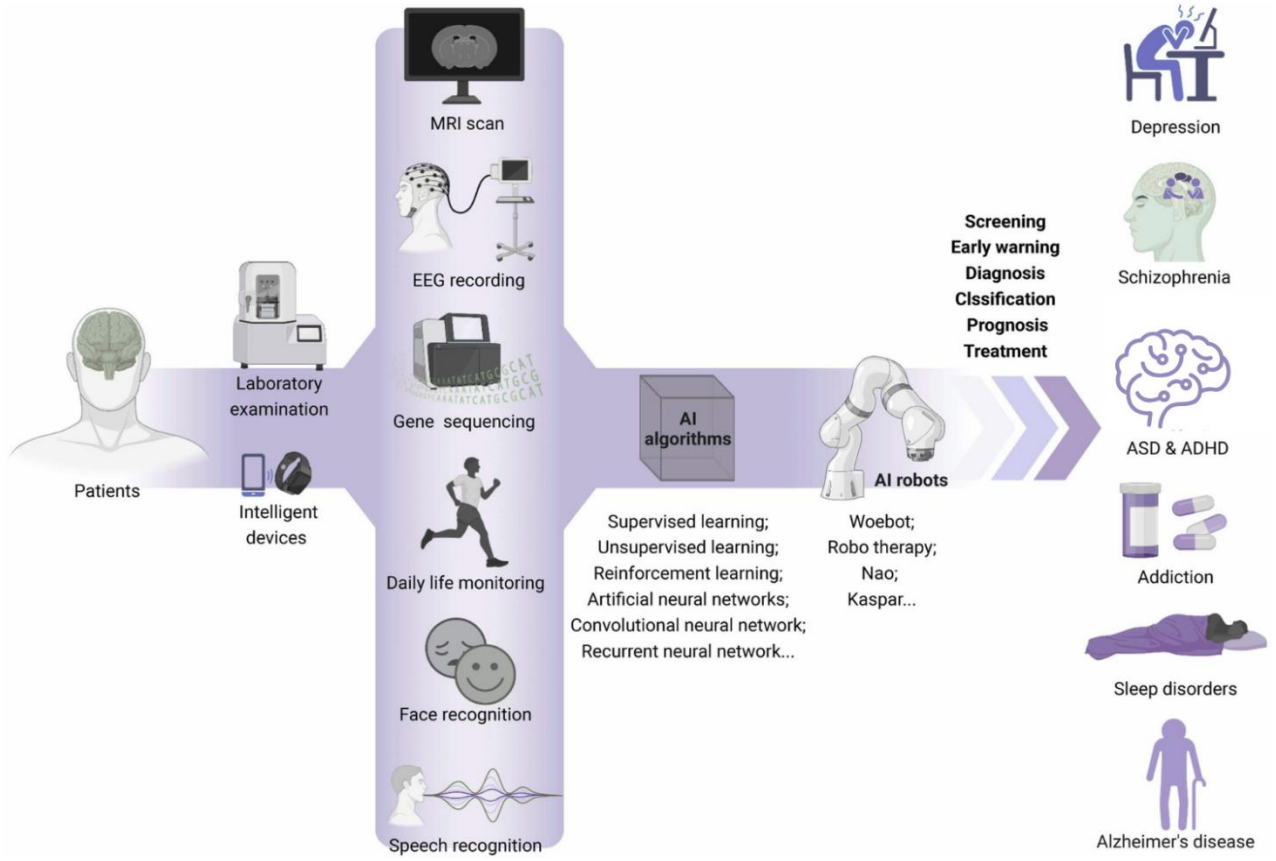

Figure 3: The potential for integrating AI technologies and multimodal data into psychiatry, extracted from Sun et al. (2023)[205]

Table 4 Examples of tools to support clinical decisions

Provider solution	Examples of trial/use by UK healthcare providers	Examples of empirical evidence
Limbic	A chatbot streamlining mental health referral processes. It was the first UKCA class IIa medical device certified tool for its functionality for mental health diagnosis support.[206,207] Examples of evidence[31,34,89,137,139,206–208]	Everyturn Mental Health, Essex Partnership University NHS Foundation Trust, Surrey and Borders Partnership NHS Foundation Trust, Nottinghamshire Healthcare NHS Foundation Trust, Lincolnshire Partnership NHS Foundation Trust, Living Well Consortium UK, Mental Health Matters. Note that they also offer Limbic Care, CBT-based[z] provision directly supporting service users.[79]

[z] Cognitive Behavioural Therapy (CBT) is a talking therapy which supports people to change the way they think and behave.

Predictix: identifying potential service user responses to treatment	Mainly trialling in Europe. Feasibility research has just been completed at Rotherham Doncaster and South Humber NHS Foundation Trust[209]	Predictix founders published several studies demonstrating feasibility of applying ML to predict service user responses to depression treatment.[210,211] A small-scale (n=19) feasibility evaluation in the UK found Predictix recommendations were followed by GPs in 86% of cases, and its use was rapid (~20 minutes).[209,212] However, the authors recommend identifying how and where Predictix fits into clinical pathways, and that thorough clinical effectiveness and health economic evaluations are needed.[212]
Psyrin: use of voice biomarkers to detect risks	No evidence of current UK trials was found.	In an academic study of 1140 individuals Psyrin demonstrated feasibility of screening for schizophrenia, bipolar disorder and major depressive disorder using voice biomarkers from speech alongside demographic information.[213,214] One of their stated aims is to support GPs to diagnose people.[214] (Note: this trial did not involve NHS patients, and was delivered by Psyrin founders with others.)

Table 5 International examples of tools being used and trialled

Provider solution	Country	Examples of evidence
The Trevor project: provides suicide prevention and crisis intervention for LGBTQ+ young people, with trained counsellors available 24/7	United States	In 2021 Trevor Project launched 'Crisis Contact Simulators' which are AI personas used to support counsellor training,[215] alongside an AI-powered crisis risk assessment tool.[216] They estimate this has supported them to train 1000+ additional volunteers.[217]
CLOVA CareCall: A public health intervention utilising advanced ML models to provide social care to those living alone (launched in 2022)	South Korea	Research in 2023 highlighted benefits of GenAI models[aa] for delivering automated check-in calls at scale to those identified as at risk of social isolation, loneliness and related health issues,[218] including: • automated risk detection through analysis of chat interactions • proactive identification of challenges, enabling escalation to human intervention where necessary but with less human capacity needed

[aa] In this case those used were specifically 'Large Language Models' (LLMs), which are AI models trained on vast amounts of data (Foundation Models) and which leverage advancements in 'Natural Language Processing' (NLP) so they can receive inputs and create outputs in natural written or spoken language. Advancements in NLP have also contributed to opportunities.

- mitigating loneliness and the emotional burden of those contacted; some describe looking forward to the check-in

Challenges included the GenAI chatbot agent giving inappropriate responses such as making false promises to meet people.[218] The developers are attempting to mitigate this.[218] (PN738)

Using Generative AI to support delivery

The recent development of GenAI has created new possibilities and benefits to enhance digital mental healthcare delivery including:

- enhancing capabilities for early detection and monitoring of mental ill-health
- creating personalised and emotional chatbot responses during delivery of mental health and wellbeing interventions; and supporting delivery of diverse therapy types[61,219,220]
- augmenting therapist work,[219,221] for example by engaging clients, in between sessions or therapists integrating AI-generated images, text, or audio into online or in-person sessions (such as when delivering writing therapy[bb] or other types of arts-based therapies[cc])[61]
- increasing the efficacy of interventions to reduce symptoms[81]
- enhancing synthetic data[dd] creation[61,220], such as model data which can be used to mitigate bias in real data

One example is GenAI enhancing avatar therapy[ee] which could benefit people experiencing visual or auditory hallucations.[223,227–229] Although a 2020 review found

[bb] Writing therapy involves a professional providing specific prompts to exercises to guide service users to use writing to analyse thoughts, feelings, or events.

[cc] Arts-based therapy can use any type of creative activity or media, and involves the therapist supporting the service user to understand more about themselves and their experiences and feelings through engaging with the creative processes.

[dd] "Synthetic data is data that has been generated using a purpose-built mathematical model or algorithm, with the aim of solving a (set of) data science task(s)."[222] It is proposed as a possible solution to privacy challenges and to support tackling bias in unrepresentative model training data.[222]

[ee] "Avatars are digital animations or self-representations that allow users to interact in a virtual environment... Furthermore, they can be personalized so that they are more realistic and are as similar as possible to the patient using them, which helps patients to better identify with them"[223] SPARX is an example of an avatar and gaming-based e-therapy program for young teenagers with depression. It has been successfully delivered since 2009 in New Zealand,[224] with measurable improvements in depression symptoms.[225,226] It is intended as treatment for mild to moderate challenges, not as a replacement for therapy.[58]

outcomes inconsistent and state longer, larger-scale and independent studies are needed.[230]

It's also suggested GenAI could support delivery of exposure therapy[ff] and that interacting with AI generated characters could offer deeper personalisation with benefits for the delivery of wellbeing interventions or peer support.[44]

Key considerations for delivery of AI tools in mental healthcare

There are additional ethical, social, and legal responsibilities to be considered, these are discussed in PN738.

Infrastructure and strategies to integrate AI across healthcare operations

Stakeholders suggest the need to improve infrastructure for successful implementation, which requires sufficient funding.[14,231] This includes data infrastructure.[16,232,233] They also propose creating a cohesive strategy to ensure the benefits of AI are realised across the NHS, rather than in isolated geographic areas where inconsistent decisions have been taken to procure AI systems (See tools being trialled across different NHS Trusts in Tables 3-5).[234] Additionally, stakeholders highlight the need for public-private partnership or public sector alternatives as the majority of trained datasets for generative AI systems are owned by private industry owners, and access to training such data can be very expensive.[235]

Designing AI systems and services

To optimise the benefits from AI systems, cross-sector stakeholders recommend ensuring clarity on where AI-technologies are appropriate,[38,236] and considering usability and different user populations.[36,80,151,237,238] For example, generalised tools could be inappropriate for military personnel who face unique stressors[239], and use specific types of language.[41] Industry and regulators also highlight possibilities to design AI tools for more ethical outcomes, such as the principle of 'Privacy by Design' which encourages developers to build in data protection measures from development stages.[240,241] See 'Designing Ethics into AI Systems' in PN738.

Stakeholders suggest revisiting service design to ensure mental healthcare services are delivered properly while leveraging AI's capabilities, and potentially redesigning services.[24 242,243] Although other stakeholders highlight the importance of effectively integrating AI within existing healthcare systems and clinical workflows.[36,80,151,237,238]

Upskilling staff

The need to upskill the UK workforce in digital technology is nationally recognised, and actions are being taken (PN637).[244] However, the Health and Social Care Committee identified challenges in digital transformation across the entire NHS

[ff] Exposure therapy can support people to reduce anxiety or fear of objects, activities or situations.

workforce including the need to upskill staff and difficulty competing with the private sector to fill specialist roles.[245] This may also relate to the lack of data science skills across the UK (PN697).

The Topol review (2019) and NHS reports suggest staff should fully understand AI issues such as validity, accuracy,[246] and building 'appropriate confidence'.[247]

Digital Fellowships upskilling staff to lead digital innovation are now provided to small cohorts of clinical and non-clinical staff working in the NHS or Social Care in England.[248] In 2023 the NHS published an AI capability framework to support building digital skills across its workforce.[249]

Evidence on cost-effectiveness

Academic and government reviews conclude there is limited evidence examining cost-effectiveness or economic evaluation.[221,250–253] Stakeholders describe risks of ensuring long-term financial sustainability,[19] highlighting costs of maintenance, updates and staff training.[253–255]

Recent research by DMHI providers includes:

- Modelling the healthcare costs of 27,540 service users with mood and anxiety disorders suggested CBT delivered online[gg] is cost-effective as service users access treatment sooner (reducing pressure on other medical services) and complete the treatment more quickly with similar effectiveness.[256]

- Cost-effectiveness of a referral tool[hh] for talking therapies services over four months was estimated as £188.25-£221.89 for every extra person who recovered. 58,475 service users were included in the analysis.[139]

- Analysis of a cohort of 2,160 children and young people who accessed a web-based platform combining peer support and therapy service[ii] over 12 months estimated a combined NHS and crime sector saving of £469,237.[258] This approach appears to take a broader systems view than often taken under the current NICE economic analysis guidance which specifies impacts on health and personal social services (4.2.7), although the guidance does specify that in exceptional circumstances the costs to other government agencies could be presented (4.4.22).[259]

Some have also suggested technical or workflow-related approaches to reducing cost.[260] Third sector research into lived experience perspectives found concerns that cost savings would be the primary consideration for introducing new digital

[gg] The intervention involved a 'stepped care' model in which people with milder symptoms were first signposted to a low-intensity more self-directed intervention, and those with more severe symptoms or complex needs to a higher-intensity intervention where there were more therapist-delivered online sessions.[256] The DMHI did not use AI.

[hh] This referral tool includes uses a chatbot, and now incorporates GenAI technology.

[ii] This platform incorporates AI to support safeguarding by assisting the work of moderating online content.[49,257]

technologies, alongside hopes that technologies could broaden the range of support available, rather than replace human support.[63]

Public engagement and trust

Academic and NHS stakeholders describe lack of service user and staff trust[jj] and digital disengagement as risks to implementation. Responsible delivery and public engagement was suggested as a mitigation (PN725).[19,62,130,263–265]

A 2024 survey found that although the majority were supportive of AI use in healthcare, 1 in 6 members of the public (out of 7,201) and 1 in 10 NHS staff (out of 1,292) think it will make the quality of care worse.[62]

Studies indicate people are less likely to trust AI unless privacy, confidentiality and data security are efficiently addressed and communicated to them.[60,64,266]

Academic and international stakeholders suggest standardised legal frameworks to govern AI tools and systems could enhance user trust and social acceptance.[267,268]

There are existing legal frameworks that apply to AI systems (Table.3). Regulatory agencies are currently considering revisions or additional guidance for AI technologies ('Regulatory Challenges and Responses' section in PN738). Stakeholders also highlight that trust in the institutions and people deploying AI tools is important.[265,269]

Academic and clinical stakeholders also highlight that the AI systems themselves need to be trustworthy.[55,261]

The risks of staff and patients trusting or relying on AI outputs too much must also be mitigated through due diligence.[53,236,270–272]

Participation and co-design

Stakeholders emphasise the benefits of engagement or co-design of AI tools with both clinicians and people with lived experience and diverse identities,[27,55,61,63,264,273–288] during all stages of developing products, standards[55,133,273,274,279,289] and governance.[278] This includes benefits in terms of shaping the agenda and building trust.[55,63,290,55 265]

Deliberative approaches[kk] are also proposed to support responsible innovation and public acceptability of AI.[291] The Ada Lovelace Institute is developing tools and

[jj] Academic analysis highlights both trust and trustworthiness (being worthy of trust) are important, and that trust can be divided into multiple factors such as: trust in others truth claims, trust in others commitments to do what they say they will, and trust in others competence to carry out actions.[261,262] They highlight the importance of defining 'who' and 'what' is being trusted.[261]

[kk] Deliberative approaches would involve AI experts exploring and debating these issues with different types of stakeholders over a period of time.[291] "Deliberation is an approach to decision-making that allows participants to consider relevant information from multiple points of view. Deliberation enables participants to discuss the issues and options and to develop their thinking together before coming to a view, taking into account the values that inform people's opinions."[292]

resources to support public involvement in AI-related policymaking and development.[293]

Education and awareness

A 2024 study found that higher levels of AI literacy were linked to higher levels of trust, particularly in safety-critical scenarios such as healthcare.[294]

Researchers argue that if clinical practitioners have insufficient knowledge of AI systems, they cannot support service users to make adequately informed consent choices.[295,296] Researchers therefore recommend targeted education initiatives[ll].[59,294,303]

References

1. Dykxhoorn, J. *et al.* (2024). Temporal patterns in the recorded annual incidence of common mental disorders over two decades in the United Kingdom: a primary care cohort study. *Psychol. Med.*, Vol 54, 663–674.
2. NHS England Digital (2023). Mental Health of Children and Young People in England, 2023 - wave 4 follow up to the 2017 survey.
3. ONS UK Measures of National Well-being Dashboard - Office for National Statistics.
4. Kirk-Wade, E. (2024). UK disability statistics: Prevalence and life experiences.
5. Lord Darzi (2024). Independent investigation of the NHS in England. Department of Health & Social Care.
6. NHS England Digital (2024). Mental Health Services Monthly Statistics, Performance September 2024.
7. The British Medical Association (2024). Mental health pressures data analysis.
8. The British Medical Association (2024). Mental health workforce report.
9. Department for Health and Social Care Media Team (2023). How we are supporting mental health services in England – Department of Health and Social Care Media Centre.
10. NHS (2019). NHS Long Term Plan » NHS Mental Health Implementation Plan 2019/20 – 2023/24.
11. NHS England (2023). NHS Long Term Workforce Plan.
12. NHS England (2024). » 2024/25 priorities and operational planning guidance.
13. NHS England (2024). NHS Talking Therapies for anxiety and depression manual.
14. National Audit Office (NAO) (2020). Digital transformation in the NHS.
15. Department of Health and Social Care (2023). Mental health and wellbeing plan: discussion paper. HM Government.
16. Joshi, I. *et al.* (2019). Artificial Intelligence: How to get it right Putting policy into practice for safe data-driven innovation in health and care.

[ll] Some suggest the ability to ask the right critical questions is important, and adults need to 'play' with AI to understand it.[297] Many free online resources for AI education exist[297,298 299–301] as well as steps to incorporate as part of public engagement.[265] Similar suggestions have been made for other emerging technologies, such as the metaverse.[302]

17. Google Cloud Introduction to prompting Generative AI on Vertex AI. *Generative AI on Vertex AI*.
18. Darcy, A. (2023). Why Generative AI Is Not Yet Ready for Mental Healthcare. *Woebot Health*.
19. Mestre, R. *et al.* (2024). Building Responsible AI for Mental Health: Insights from the First RAI4MH Workshop. University of Southampton, and Institute for Experiential AI (Northeastern University).
20. WHO (2023). Artificial intelligence in mental health research: new WHO study on applications and challenges.
21. Rebelo, A. D. *et al.* (2023). The impact of artificial intelligence on the tasks of mental healthcare workers: A scoping review. *Comput. Hum. Behav. Artif. Hum.*, Vol 1, 100008.
22. NHS England (2023). Digitally enabled triage.
23. Iyortsuun, N. K. *et al.* (2023). A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis. *Healthcare*, Vol 11, 285.
24. Asbach, M. *et al.* (2024). AI in Psychiatry: Changing the Landscape of Mental Health Care. *Psychiatr. Times*, Vol 14, 15–17. MJH Life Sciences.
25. Cecula, P. *et al.* (2021). Applications of artificial intelligence to improve patient flow on mental health inpatient units - Narrative literature review. *Heliyon*, Vol 7,
26. Salazar de Pablo, G. *et al.* (2021). Implementing Precision Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice. *Schizophr. Bull.*, Vol 47, 284–297.
27. Khan, N. Z. *et al.* (2022). Use of Artificial Intelligence-Based Strategies for Assessing Suicidal Behavior and Mental Illness: A Literature Review. *Cureus*, Vol 14, e27225.
28. Kessler, R. C. *et al.* (2017). Predicting suicides after outpatient mental health visits in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Mol. Psychiatry*, Vol 22, 544–551. Nature Publishing Group.
29. Chekroud, A. M. *et al.* (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, Vol 20, 154–170.
30. Aafjes-van Doorn, K. *et al.* (2021). A scoping review of machine learning in psychotherapy research. *Psychother. Res.*, Vol 31, 92–116.
31. Hornstein, S. *et al.* (2023). Personalization strategies in digital mental health interventions: a systematic review and conceptual framework for depressive symptoms. *Front. Digit. Health*, Vol 5, Frontiers.
32. Deeney (2024). Can stratified care help primary care manage long-term conditions? *NIHR Evidence*.
33. Delgadillo, J. *et al.* (2022). Stratified Care vs Stepped Care for Depression: A Cluster Randomized Clinical Trial. *JAMA Psychiatry*, Vol 79, 101–108.
34. Rollwage, M. *et al.* (2023). Using Conversational AI to Facilitate Mental Health Assessments and Improve Clinical Efficiency Within Psychotherapy Services: Real-World Observational Study. *JMIR AI*, Vol 2, e44358.
35. Olawade, D. B. *et al.* (2024). Enhancing mental health with Artificial Intelligence: Current trends and future prospects. *J. Med. Surg. Public Health*, Vol 3, 100099.
36. Balcombe, L. *et al.* (2022). Human-Computer Interaction in Digital Mental Health. *Informatics*, Vol 9, 14.
37. Graham, S. *et al.* (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Curr. Psychiatry Rep.*, Vol 21, 116.
38. The British Medical Association (BMA) (2024). Principles for artificial intelligence (AI) and its application in healthcare.

39. Dawoodbhoy, F. M. *et al.* (2021). AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. *Heliyon*, Vol 7,
40. Ettman, C. K. *et al.* (2023). The Potential Influence of AI on Population Mental Health. *JMIR Ment. Health*, Vol 10, e49936.
41. Vallejos, E. P. (2024). Mindful Mortality: Digital Mental Health and the Dying Process.
42. Xygkou, A. *et al.* (2023). The 'Conversation' about Loss: Understanding How Chatbot Technology was Used in Supporting People in Grief. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15. Association for Computing Machinery.
43. Fabry, R. E. *et al.* (2024). The Affective Scaffolding of Grief in the Digital Age: The Case of Deathbots. *Topoi*, Vol 43, 757–769.
44. Pataranutaporn, P. *et al.* (2021). AI-generated characters for supporting personalized learning and well-being. *Nat. Mach. Intell.*, Vol 3, 1013–1022. Nature Publishing Group.
45. NHS England Social prescribing.
46. Kenji ARTS x WELL-BEING.
47. Nadarasa, A. (2024). Social prescribing in the metaverse: a new frontier for primary care practice. *Glob. Health J.*, Vol 8, 32–35.
48. Marshall, P. *et al.* (2024). Understanding the Impacts of Online Mental Health Peer Support Forums: Realist Synthesis. *JMIR Ment. Health*, Vol 11, e55750.
49. Bernard, A. (2024). Responsible Innovation in Practice: NLP for Safeguarding Assistance (Kooth PLC).
50. Wellcome Trust (2024). Mental Health Award: Accelerating scalable digital mental health interventions - Funding.
51. Wellcome Trust (2024). Personal Communication.
52. Wellcome Trust What we do and don't fund in Mental Health | Grant Funding.
53. NHS England (2023). Artificial intelligence (AI) and machine learning.
54. Humphreys, J. *et al.* (2024). *Digital Mental Health Technology: User and Public Perspectives*. Medicines and Healthcare products Regulatory Agency.
55. Muchamore, I. *et al.* (2024). How lived experience expertise shapes research and development in digital mental health: A Review of literature and Insights. The Wellcome Trust.
56. Lederman, R. *et al.* (2021). The Digital Therapeutic Alliance: Prospects and Considerations. *JMIR Ment. Health*, Vol 8, e31385.
57. Palmer, C. E. *et al.* (2024). Combining AI and human support in mental health: a digital intervention with comparable effectiveness to human-delivered care. 2024.07.17.24310551. medRxiv.
58. Jackson (2023). The big issue: The brave new world of AI therapy. *Therapy Today*.
59. De Choudhury, M. *et al.* (2023). Benefits and Harms of Large Language Models in Digital Mental Health. arXiv.
60. Gentry, E. *et al.* (2024). 'I don't trust it, but I have to trust it': The Paradox of Trust vs Use of Online Technology Across The Mental Health Spectrum. OSF.
61. Xian, X. *et al.* (2024). Debate and Dilemmas Regarding Generative AI in Mental Health Care: Scoping Review. *Interact. J. Med. Res.*, Vol 13, e53672.
62. Thornton, N. *et al.* (2024). AI in health care: what do the public and NHS staff think? The Health Foundation.
63. Rethink Mental Illness (2022). Summary paper: Engaging experts by experience about the role of digital technology in the future of mental health care.
64. Chapman, A. *et al.* (2024). Sociotechnical Considerations for Accessibility and Equity in AI for

Healthcare. in *Companion Proceedings of the ACM Web Conference 2024*. 1158–1161. ACM.
65. D'Alfonso, S. (2020). AI in mental health. *Curr. Opin. Psychol.*, Vol 36, 112–117.
66. Macrae, C. (2024). Managing risk and resilience in autonomous and intelligent systems: Exploring safety in the development, deployment, and use of artificial intelligence in healthcare. *Risk Anal.*, Vol n/a,
67. Thirunavukarasu, A. J. (2023). Large language models will not replace healthcare professionals: curbing popular fears and hype. *J. R. Soc. Med.*, Vol 116, 181–182. SAGE Publications.
68. Ly, K. H. *et al.* (2017). A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interv.*, Vol 10, 39–46.
69. Fitzpatrick, K. K. *et al.* (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment. Health*, Vol 4, e19.
70. Lin, I. W. *et al.* (2024). IMBUE: Improving Interpersonal Effectiveness through Simulation and Just-in-time Feedback with Human-Language Model Interaction. arXiv.
71. Sharma, A. *et al.* (2023). Cognitive Reframing of Negative Thoughts through Human-Language Model Interaction. arXiv.
72. Inkster, B. *et al.* (2018). An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR MHealth UHealth*, Vol 6, e12106.
73. Eis, S. *et al.* (2022). Mobile Applications in Mood Disorders and Mental Health: Systematic Search in Apple App Store and Google Play Store and Review of the Literature. *Int. J. Environ. Res. Public. Health*, Vol 19, 2186. Multidisciplinary Digital Publishing Institute.
74. Larsen, M. E. *et al.* (2019). Using science to sell apps: Evaluation of mental health app store quality claims. *Npj Digit. Med.*, Vol 2, 1–6. Nature Publishing Group.
75. Simon, D. A. *et al.* (2022). Skating the line between general wellness products and regulated devices: strategies and implications. *J. Law Biosci.*, Vol 9, lsac015.
76. Kretzschmar, K. *et al.* (2019). Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support. *Biomed. Inform. Insights*, Vol 11, 1178222619829083. SAGE Publications Ltd STM.
77. Information Commissioner's Office (2024). Rights related to automated decision making including profiling. *Guidance*. ICO.
78. Adamopoulou, E. *et al.* (2020). Chatbots: History, technology, and applications. *Mach. Learn. Appl.*, Vol 2, 100006.
79. McFadyen, J. *et al.* (2024). AI-enabled conversational agent increases engagement with cognitive-behavioral therapy: A randomized controlled trial. 2024.11.01.24316565. medRxiv.
80. Casu, M. *et al.* (2024). AI Chatbots for Mental Health: A Scoping Review of Effectiveness, Feasibility, and Applications. *Appl. Sci.*, Vol 14, 5889. Multidisciplinary Digital Publishing Institute.
81. Li, H. *et al.* (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *Npj Digit. Med.*, Vol 6, 1–14.
82. Fulmer, R. *et al.* (2018). Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. *JMIR Ment. Health*, Vol 5, e64.

83. Pham, K. T. *et al.* (2022). Artificial Intelligence and Chatbots in Psychiatry. *Psychiatr. Q.*, Vol 93, 249–253.
84. Beg, M. J. *et al.* (2024). Artificial Intelligence for Psychotherapy: A Review of the Current State and Future Directions. *Indian J. Psychol. Med.*, 02537176241260819. SAGE Publications India.
85. He, Y. *et al.* (2023). Conversational Agent Interventions for Mental Health Problems: Systematic Review and Meta-analysis of Randomized Controlled Trials. *J. Med. Internet Res.*, Vol 25, e43862.
86. Zhong, W. *et al.* (2024). The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. *J. Affect. Disord.*, Vol 356, 459–469.
87. Maia, E. *et al.* (2023). Empowering Preventive Care with GECA Chatbot. *Healthcare*, Vol 11, 2532. Multidisciplinary Digital Publishing Institute.
88. Habicht, J. *et al.* (2024). Generative AI-Enabled Therapy Support Tool Improves Clinical Outcomes and Patient Engagement in NHS Talking Therapies. OSF.
89. Habicht, J. *et al.* (2024). Closing the accessibility gap to mental health treatment with a conversational AI-enabled self-referral tool. *Nat. Med.*, 595–602.
90. Tong, F. *et al.* (2023). Conceptualizing the digital therapeutic alliance in the context of fully automated mental health apps: A thematic analysis. *Clin. Psychol. Psychother.*, Vol 30, 998–1012.
91. MacNeill, A. L. *et al.* (2024). Effectiveness of a Mental Health Chatbot for People With Chronic Diseases: Randomized Controlled Trial. *JMIR Form. Res.*, Vol 8, e50025.
92. Brooks, R. (2023). I tried the Replika AI companion and can see why users are falling hard. The app raises serious ethical questions. *The Conversation*.
93. Mozilla Foundation (2024). Creepy.exe: Mozilla Urges Public to Swipe Left on Romantic AI Chatbots Due to Major Privacy Red Flags.
94. Skjuve, M. *et al.* (2021). My Chatbot Companion - a Study of Human-Chatbot Relationships. *Int. J. Hum.-Comput. Stud.*, Vol 149, 102601.
95. Ta, V. *et al.* (2020). User Experiences of Social Support From Companion Chatbots in Everyday Contexts: Thematic Analysis. *J. Med. Internet Res.*, Vol 22, e16235.
96. VoiceBox (2023). Coded Companions: young people's relationships with AI chatbots.
97. Crawford, J. *et al.* (2024). 1 in 3 people are lonely. Will AI help, or make things worse? *The Conversation*.
98. Ma, Z. *et al.* (2024). Understanding the Benefits and Challenges of Using Large Language Model-based Conversational Agents for Mental Well-being Support. *AMIA. Annu. Symp. Proc.*, Vol 2023, 1105–1114.
99. Maples, B. *et al.* (2024). Loneliness and suicide mitigation for students using GPT3-enabled chatbots. *Npj Ment. Health Res.*, Vol 3, 1–6. Nature Publishing Group.
100. Laestadius, L. *et al.* (2024). Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika. *New Media Soc.*, Vol 26, 5923–5941. SAGE Publications.
101. Xie, T. *et al.* (2022). Attachment Theory as a Framework to Understand Relationships with Social Chatbots: A Case Study of Replika.
102. Xiang, C. (2023). 'He Would Still Be Here': Man Dies by Suicide After Talking with AI Chatbot, Widow Says. *VICE*.

103. King, J. (2024). Boy, 14, killed himself after 'falling in love' with AI chatbot. *Metro*.
104. Sharma, A. *et al.* (2023). Cognitive Reframing of Negative Thoughts through Human-Language Model Interaction. arXiv.
105. National Institute for Health and Care Excellence Technology appraisal guidance. *NICE*. NICE.
106. National Institute for Health and Care Excellence Early Value Assessment (EVA) for medtech. *What we do*. NICE.
107. National Institute for Health and Care Excellence (2023). The technology | Digitally enabled therapies for adults with anxiety disorders: early value assessment | |. NICE.
108. National Institute for Health and Care Excellence (2023). The technology | Digitally enabled therapies for adults with depression: early value assessment. NICE.
109. National Institute for Health and Care Excellence (2024). The technologies | Digital health technologies to help manage symptoms of psychosis and prevent relapse in adults and young people: early value assessment. NICE.
110. National Institute for Health and Care Excellence (2023). Recommendations | Digitally enabled therapies for adults with anxiety disorders: early value assessment. NICE.
111. National Institute for Health and Care Excellence (2023). Recommendations | Digitally enabled therapies for adults with depression: early value assessment. NICE.
112. National Institute for Health and Care Excellence (2024). Recommendations | Digital health technologies to help manage symptoms of psychosis and prevent relapse in adults and young people: early value assessment. NICE.
113. Grodniewicz, J. P. *et al.* (2023). Waiting for a digital therapist: three challenges on the path to psychotherapy delivered by artificial intelligence. *Front. Psychiatry*, Vol 14,
114. Brown, J. E. H. *et al.* (2021). AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare. *SSM - Ment. Health*, Vol 1, 100017.
115. Wysa - Everyday Mental Health (2022). 'All Worked Up' UK edition.
116. Leo, A. J. *et al.* (2022). A Digital Mental Health Intervention in an Orthopedic Setting for Patients With Symptoms of Depression and/or Anxiety: Feasibility Prospective Cohort Study. *JMIR Form. Res.*, Vol 6, e34889.
117. Leo, A. J. *et al.* (2022). Digital Mental Health Intervention Plus Usual Care Compared With Usual Care Only and Usual Care Plus In-Person Psychological Counseling for Orthopedic Patients With Symptoms of Depression or Anxiety: Cohort Study. *JMIR Form. Res.*, Vol 6, e36203.
118. Farrand, P. *et al.* (2024). Use and Engagement With Low-Intensity Cognitive Behavioral Therapy Techniques Used Within an App to Support Worry Management: Content Analysis of Log Data. *JMIR MHealth UHealth*, Vol 12, e47321.
119. (2024). Iona Mind. *DigitalHealth.London*.
120. Lewis, S. *et al.* (2020). Smartphone-Enhanced Symptom Management In Psychosis: Open, Randomized Controlled Trial. *J. Med. Internet Res.*, Vol 22, e17019.
121. Lewis, S. *et al.* (2018). O11.1. A RANDOMISED CONTROLLED TRIAL OF SMARTPHONE ACTIVE SYMPTOM MONITORING IN PSYCHOSIS. *Schizophr. Bull.*, Vol 44, S106.
122. Hassan, L. *et al.* (2023). User engagement in a randomised controlled trial for a digital health intervention for early psychosis (Actissist 2.0 trial). *Psychiatry Res.*, Vol 329, 115536.

123. Bucci, S. *et al.* (2024). Effects of Actissist, a digital health intervention for early psychosis: A randomized clinical trial. *Psychiatry Res.*, Vol 339, 116025.
124. Ewbank, M. P. *et al.* (2021). Understanding the relationship between patient language and outcomes in internet-enabled cognitive behavioural therapy: A deep learning approach to automatic coding of session transcripts. *Psychother. Res.*, Vol 31, 300–312. Routledge.
125. Palmer, C. E. *et al.* (2024). Combining AI and human support in mental health: a digital intervention with comparable effectiveness to human-delivered care. 2024.07.17.24310551. medRxiv.
126. Stevens, M. *et al.* (2022). Pilot evaluation to assess the effectiveness of youth peer community support via the Kooth online mental wellbeing website. *BMC Public Health*, Vol 22, 1903.
127. Digital Mental Health Care. *Kooth plc*.
128. Banwell, E. *et al.* (2022). The Helpfulness of Web-Based Mental Health and Well-being Forums for Providing Peer Support for Young People: Cross-sectional Exploration. *JMIR Form. Res.*, Vol 6, e36432. JMIR Publications.
129. World Health Organisation (2024). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models.
130. Stade, E. C. *et al.* (2024). Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation. *NPJ Ment. Health Res.*, Vol 3, 12.
131. Shatte, A. B. R. *et al.* (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychol. Med.*, Vol 49, 1426–1448.
132. Ryan, C. (2024). Feature: Into the loop – AI in mental health. *MIP - The union for health and care managers*.
133. Lawrence, H. R. *et al.* (2024). The Opportunities and Risks of Large Language Models in Mental Health. *JMIR Ment. Health*, Vol 11, e59479–e59479.
134. NHS England NHS Talking Therapies, for anxiety and depression.
135. NHS England NHS Talking Therapies, for anxiety and depression.
136. AI and Digital Regulations Service for Health and Social Care (2024). Using an AI Chatbot to ease staff shortages and improve patient care. *AI and Digital Regulations Service for health and social care*.
137. NHS Transformation Directorate (2022). Using an AI chatbot to streamline mental health referrals.
138. NHS Digital (2021). Outcomes.
139. Rollwage, M. *et al.* (2024). Conversational AI facilitates mental health assessments and is associated with improved recovery rates. *BMJ Innov.*, Vol 10, BMJ Specialist Journals.
140. Health Innovation Network (2023). Preventing clinician burnout: could Ambient Voice Technology (AVT) be key?
141. Han, R. *et al.* (2024). Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. *Lancet Digit. Health*, Vol 6, e367–e373. Elsevier.
142. Tomasik, J. *et al.* (2024). Metabolomic Biomarker Signatures for Bipolar and Unipolar Depression. *JAMA Psychiatry*, Vol 81, 101–106.
143. Tomasik, J. *et al.* (2021). A machine learning algorithm to differentiate bipolar disorder from major depressive disorder using an online mental health questionnaire and blood biomarker data. *Transl. Psychiatry*, Vol 11, 1–12. Nature Publishing Group.
144. Martin-Key, N. A. *et al.* (2021). Toward an Extended Definition of Major Depressive Disorder

Symptomatology: Digital Assessment and Cross-validation Study. *JMIR Form. Res.*, Vol 5, e27908.
145. Psychology Today Precision Psychiatry.
146. Dellen, E. van (2024). Precision psychiatry: predicting predictability. *Psychol. Med.*, Vol 54, 1500–1509.
147. Hewis, J. (2023). A salutogenic approach: Changing the paradigm. *J. Med. Imaging Radiat. Sci.*, Vol 54, S17–S21. Elsevier.
148. Saraceni, G. (2023). ARTIFICIAL INTELLIGENCE AND MENTAL HEALTH. *Humanit. RIGHTS Glob. Netw. J.*, Vol 5,
149. Mestre, R. (2024). Personal Communication.
150. Kumar, S. *et al.* (2024). Resp-BoostNet: Mental Stress Detection from Biomarkers Measurable by Smartwatches using Boosting Neural Network Technique. *IEEE Access*, 1–1.
151. Torous, J. *et al.* (2021). The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*, Vol 20, 318–335.
152. Liakata, M. (2020). Language technology and AI can aid mental health during COVID-19 and beyond. *The Alan Turing Institute*.
153. Zhang, T. *et al.* (2022). Natural language processing applied to mental illness detection: a narrative review. *Npj Digit. Med.*, Vol 5, 1–13.
154. Milne-Ives, M. *et al.* (2022). Artificial intelligence and machine learning in mobile apps for mental health: A scoping review. *PLOS Digit. Health*, Vol 1, e0000079.
155. Sinvani, R.-T. *et al.* (2021). Ecological Momentary Mood, Resilience, and Mental Health Status as Predictors of Quality of Life Among Young Adults Under Stress: A Structural Equation Modeling Analysis. *Front. Psychiatry*, Vol 12, Frontiers.
156. Koutsouleris, N. *et al.* (2016). Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry*, Vol 3, 935–946.
157. Goodday, S. M. *et al.* (2020). The True Colours Remote Symptom Monitoring System: A Decade of Evolution. *J. Med. Internet Res.*, Vol 22, e15188.
158. De Choudhury, M. *et al.* (2016). Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2098–2110. Association for Computing Machinery.
159. Huckins, J. F. *et al.* (2019). Fusing Mobile Phone Sensing and Brain Imaging to Assess Depression in College Students. *Front. Neurosci.*, Vol 13, 248.
160. Melcher, J. *et al.* (2020). Digital phenotyping for mental health of college students: a clinical review. *BMJ Ment Health*, Vol 23, 161–166. Royal College of Psychiatrists.
161. Faurholt-Jepsen, M. *et al.* (2019). Smartphone-based self-monitoring in bipolar disorder: evaluation of usability and feasibility of two systems. *Int. J. Bipolar Disord.*, Vol 7, 1.
162. Broulidakis, M. J. *et al.* (2023). Optomyography-based sensing of facial expression derived arousal and valence in adults with depression. *Front. Psychiatry*, Vol 14,
163. Abdullah, S. *et al.* (2018). Sensing Technologies for Monitoring Serious Mental Illnesses. *IEEE Multimed.*, Vol 25, 61–75.
164. Grünerbl, A. *et al.* (2015). Smartphone-Based Recognition of States and State Changes in Bipolar Disorder Patients. *IEEE J. Biomed. Health Inform.*, Vol 19, 140–148.
165. Gruenerbl, A. *et al.* (2014). Using smart phone mobility traces for the diagnosis of depressive and manic episodes in bipolar patients. in *Proceedings of the 5th Augmented Human International Conference*. 1–8. Association for Computing Machinery.

166. Gravenhorst, F. *et al.* (2015). Mobile phones as medical devices in mental disorder treatment: an overview. *Pers. Ubiquitous Comput.*, Vol 19, 335–353.
167. Melcher, J. *et al.* (2020). Digital phenotyping for mental health of college students: a clinical review. *BMJ Ment Health*, Vol 23, 161–166. Royal College of Psychiatrists.
168. Ahmed, A. *et al.* (2023). Wearable devices for anxiety & depression: A scoping review. *Comput. Methods Programs Biomed. Update*, Vol 3, 100095.
169. Use Apple Watch to practice mindfulness. *Apple Support*.
170. Udavant, S. (2022). The best mental health wearables to help with stress. *Digital Trends*.
171. Pardamean, B. *et al.* (2020). Quantified Self-Using Consumer Wearable Device: Predicting Physical and Mental Health. *Healthc. Inform. Res.*, Vol 26, 83–92. Korean Society of Medical Informatics.
172. Kaplan, D. M. *et al.* (2023). Wear With Care: A Call for Empirical Investigations of Adverse Outcomes of Consumer Health Wearables. *Mayo Clin. Proc. Digit. Health*, Vol 1, 413–418. Elsevier.
173. Hickey, B. A. *et al.* (2021). Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions and Stress: A Systematic Review. *Sensors*, Vol 21, 3461. Multidisciplinary Digital Publishing Institute.
174. D'Arcy, M. and M. (2023). What could AI mean for our money and mental health? *Money and Mental Health Policy Institute*.
175. Money and Mental Health (2023). Policy response on artificial intelligence and machine learning. *Money and Mental Health Policy Institute*.
176. Inkster, B. *et al.* (2019). Improving insights into health care with data linkage to financial technology. *Lancet Digit. Health*, Vol 1, e110–e112. Elsevier.
177. Erickson, C. M. *et al.* (2024). Alzheimer's in the modern age: Ethical challenges in the use of digital monitoring to identify cognitive changes. *Inform. Health Soc. Care*, Vol 49, 1–13. Taylor & Francis.
178. Abbe, A. *et al.* (2015). Text mining applications in psychiatry: a systematic literature review. *Int. J. Methods Psychiatr. Res.*, Vol 25, 86.
179. Calvo, R. A. *et al.* (2017). Natural language processing in mental health applications using non-clinical texts. *Nat. Lang. Eng.*, Vol 23, 649–685.
180. Montejo-Ráez, A. *et al.* (2024). A survey on detecting mental disorders with natural language processing: Literature review, trends and challenges. *Comput. Sci. Rev.*, Vol 53, 100654.
181. Bey, R. *et al.* (2024). Natural language processing of multi-hospital electronic health records for public health surveillance of suicidality. *Npj Ment. Health Res.*, Vol 3, 1–9. Nature Publishing Group.
182. Merhbene, G. *et al.* (2024). Investigating machine learning and natural language processing techniques applied for detecting eating disorders: a systematic literature review. *Front. Psychiatry*, Vol 15, Frontiers.
183. Guerdelli, H. *et al.* (2022). Macro- and Micro-Expressions Facial Datasets: A Survey. *Sensors*, Vol 22, 1524.
184. Ripke, S. *et al.* (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, Vol 511, 421–427. Nature Publishing Group.
185. Wray, N. R. *et al.* (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.*, Vol 50, 668–681.
186. Newby, D. *et al.* (2024). Optimising the use of electronic

medical records for large scale research in psychiatry. *Transl. Psychiatry*, Vol 14, 1–10. Nature Publishing Group.
187. Oliver, D. *et al.* (2021). Real-world implementation of precision psychiatry: Transdiagnostic risk calculator for the automatic detection of individuals at-risk of psychosis. *Schizophr. Res.*, Vol 227, 52–60.
188. Low, D. M. *et al.* (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investig. Otolaryngol.*, Vol 5, 96–116.
189. Cummins, N. *et al.* (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Commun.*, Vol 71, 10–49.
190. Vasudevan, S. *et al.* (2022). Digital biomarkers: Convergence of digital health technologies and biomarkers. *Npj Digit. Med.*, Vol 5, 1–3. Nature Publishing Group.
191. Lalousis, P. A. *et al.* (2024). Anhedonia as a Potential Transdiagnostic Phenotype With Immune-Related Changes in Recent-Onset Mental Health Disorders. *Biol. Psychiatry*, Vol 96, 615–622. Elsevier.
192. Lalousis, P. A. *et al.* (2022). Neurobiologically Based Stratification of Recent-Onset Depression and Psychosis: Identification of Two Distinct Transdiagnostic Phenotypes. *Biol. Psychiatry*, Vol 92, 552–562. Elsevier.
193. Bzdok, D. *et al.* (2018). Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging*, Vol 3, 223–230.
194. Pigoni, A. *et al.* (2024). Machine learning and the prediction of suicide in psychiatric populations: a systematic review. *Transl. Psychiatry*, Vol 14, 1–22.
195. Sajjadian, M. *et al.* (2021). Machine learning in the prediction of depression treatment outcomes: a systematic review and meta-analysis. *Psychol. Med.*, Vol 51, 2742–2751.
196. Meehan, A. J. *et al.* (2022). Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges. *Mol. Psychiatry*, Vol 27, 2700–2708.
197. Burr, C. *et al.* (2020). Digital Psychiatry: Risks and Opportunities for Public Health and Wellbeing. *IEEE Trans. Technol. Soc.*, Vol 1, 21–33.
198. Public Health England (2018). Chapter 6: wider determinants of health. *GOV.UK*.
199. Worley, R. *et al.* (2022). 'Artificial intelligence and public (dis-)trust? A case study on using artificial intelligence in mental health services', funded by the British Academy and Leverhulme Trust: SG2122\210037. *[Results pending publication]*.
200. Addelman, M. (2022). Digital psychosis monitoring system trial launches.
201. CONNECT Digital Study.
202. Sheffield Health and Social Care (2024). New research to transform care and outcomes for people with personality disorders in Sheffield.
203. UKRI (2024). Accelerating research in severe mental illness.
204. CARE For Researcher.
205. Sun, J. *et al.* (2023). Artificial intelligence in psychiatry research, diagnosis, and therapy. *Asian J. Psychiatry*, Vol 87, 103705.
206. Carrington, B. (2023). AI mental health chatbot that predicts disorders becomes first in world to gain Class IIa UKCA medical device status. *Limbic / Newa*.
207. Longworth, C. (2023). Talk to the bot: AI assistant marks breakthrough for UK mental health. *Medical Device Network*.
208. Health Research Authority (2022). Limbic: AI-Enabled Information Collection Tool for Mental Health.
209. Health Research Authority (2023). Predictix. *Planning & improving research*.
210. Taliaz, D. *et al.* (2021). Optimizing prediction of response to

antidepressant medications using machine learning and integrated genetic, clinical, and demographic data. *Transl. Psychiatry*, Vol 11, 1–9. Nature Publishing Group.
211. Spinrad, A. *et al.* (2021). Optimizing prediction of response to antidepressant medications using machine learning and environmental data. *Eur. Psychiatry*, Vol 64, S755–S755.
212. Kellett, S. (2024). Results From a Feasibility Field Trial of the Integration of Predictix for the Management and Recognition of Depression in General Practice.
213. Olah, J. *et al.* (2024). Detecting schizophrenia, bipolar disorder, psychosis vulnerability and major depressive disorder from 5 minutes of online-collected speech.
214. Wong, W. L. E. (2024). Artificial Intelligence (AI) Powered Psychosis Assessment Tool for Primary Care. *NIHR Funding and Awards*.
215. Trevor News (2021). The Trevor Project Expands Its AI Innovation with New "Crisis Contact Simulator" Persona to Scale Counselor Training. *The Trevor Project*.
216. Trevor News (2021). The Trevor Project Launches New AI Tool To Support Crisis Counselor Training. *The Trevor Project*.
217. Kerrigan, P. *et al.* (2023). Automating vulnerability: Algorithms, artificial intelligence and machine learning for gender and sexual minorities. in *Routledge Handbook of Sexuality, Gender, Health and Rights*. Routledge.
218. Jo, E. *et al.* (2023). Understanding the Benefits and Challenges of Deploying Conversational AI Leveraging Large Language Models for Public Health Intervention. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16. ACM.
219. Siddals, S. *et al.* (2024). "It happened to be the perfect thing": experiences of generative AI chatbots for mental health. *Npj Ment. Health Res.*, Vol 3, 1–9. Nature Publishing Group.
220. Sai, S. *et al.* (2024). Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies, and Limitations. *IEEE Access*, Vol 12, 31078–31106.
221. Sezgin, E. *et al.* (2024). Behavioral health and generative AI: a perspective on future of therapies and patient care. *Npj Ment. Health Res.*, Vol 3, 1–6. Nature Publishing Group.
222. Jordon, J. *et al.* (2022). Synthetic Data -- what, why and how? arXiv.
223. Franco, M. *et al.* (2021). Could Avatar Therapy Enhance Mental Health in Chronic Patients? A Systematic Review. *Electronics*, Vol 10, 2212. Multidisciplinary Digital Publishing Institute.
224. Fleming, T. *et al.* (2021). Technology Matters: SPARX – computerised cognitive behavioural therapy for adolescent depression in a game format. *Child Adolesc. Ment. Health*, Vol 26, 92–94. John Wiley & Sons, Ltd.
225. Fleming, T. *et al.* (2012). A Pragmatic Randomized Controlled Trial of Computerized CBT (SPARX) for Symptoms of Depression among Adolescents Excluded from Mainstream Education. *Behav. Cogn. Psychother.*, Vol 40, 529–541.
226. Merry, S. N. *et al.* (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *BMJ*, Vol 344, e2598. British Medical Journal Publishing Group.
227. Ward, T. *et al.* (2020). AVATAR Therapy for Distressing Voices: A Comprehensive Account of Therapeutic Targets. *Schizophr. Bull.*, Vol 46, 1038.

228. Craig, T. K. *et al.* (2018). AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. *Lancet Psychiatry*, Vol 5, 31–40. Elsevier.
229. Garety, P. A. Digital AVATAR therapy for distressing voices in psychosis: the phase 2/3 AVATAR2 trial. *Nat. Med.*,
230. Aali, G. *et al.* (2020). Avatar Therapy for people with schizophrenia or related disorders. *Cochrane Database Syst. Rev.*, John Wiley & Sons, Ltd.
231. Karpathakis, K. *et al.* (2024). A Justifiable Investment in AI for Healthcare: Aligning Ambition with Reality. *Minds Mach.*, Vol 34, 38.
232. Science, Innovation and Technology Committee (2024). *Legacy – Parliament 2019–24*. House of Commons.
233. NHS (2019). The NHS Long Term Plan.
234. Thronton, N. *et al.* (2024). Priorities for an AI in health care strategy. The Health Foundation.
235. World Economic Forum (2024). How public-private partnerships can ensure AI for good. *World Economic Forum*.
236. Fusar-Poli, P. *et al.* (2022). Ethical considerations for precision psychiatry: A roadmap for research and clinical practice. *Eur. Neuropsychopharmacol.*, Vol 63, 17–34.
237. Abd-Alrazaq, A. A. *et al.* (2021). Perceptions and Opinions of Patients About Mental Health Chatbots: Scoping Review. *J. Med. Internet Res.*, Vol 23, e17828.
238. Milne-Ives, M. *et al.* (2020). The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review. *J. Med. Internet Res.*, Vol 22, e20346.
239. Middleton, S. *et al.* (2024). AI for Defence: Readiness, Resilience and Mental Health. *RUSI J.*, 52–62.
240. Kadi Coult Wharton The 7 principles of privacy by design.
241. Information Commissioner's Office (2024). Data protection by design and by default. ICO.
242. Downe, L. 15 Principles of Good Service Design. *Good Services*.
243. olbo.design How to understand and shape AI-enabled services and operations? *Service Design Tools*.
244. Fiske, A. *et al.* (2019). Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *J. Med. Internet Res.*, Vol 21, e13216.
245. Health and Social Care Committee (2023). Digital transformation in the NHS. House of Commons.
246. Topol, E. (2019). The Topol Review: Preparing the healthcare workforce to deliver the digital future. Department of Health and Social Care.
247. NHS England Moving from trust to appropriate confidence. *NHS England | Workforce, training and education | Digital Transformation*.
248. NHS Health Education England Topol Digital Fellowships. *The Topol Review*.
249. NHS England (2023). Artificial Intelligence (AI) and Digital Healthcare Technologies Capability Framework published today. *NHS England | Workforce, training and education | Digital Transformation*.
250. Sin, J. *et al.* (2020). Digital Interventions for Screening and Treating Common Mental Disorders or Symptoms of Common Mental Illness in Adults: Systematic Review and Meta-analysis. *J. Med. Internet Res.*, Vol 22, e20581.
251. Hannah, L. A. *et al.* (2023). Economic evaluation of interventions for treatment-resistant depression: A systematic review. *Front. Psychiatry*, Vol 14,
252. Milne-Ives, M. *et al.* (2020). The Effectiveness of Artificial Intelligence Conversational Agents in Health Care:

Systematic Review. *J. Med. Internet Res.*, Vol 22, e20346.
253. Hopkin, G. *et al.* (2024). Considerations for regulation and evaluation of digital mental health technologies. *Digit. Health*, Vol 10, 20552076241293313. SAGE Publications Ltd.
254. Sutton, R. T. *et al.* (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *Npj Digit. Med.*, Vol 3, 1–10. Nature Publishing Group.
255. Bélisle-Pipon, J.-C. *et al.* (2021). What Makes Artificial Intelligence Exceptional in Health Technology Assessment? *Front. Artif. Intell.*, Vol 4, Frontiers.
256. Catarino, A. *et al.* (2023). Economic evaluation of 27,540 patients with mood and anxiety disorders and the importance of waiting time and clinical effectiveness in mental healthcare. *Nat. Ment. Health*, Vol 1, 667–678. Nature Publishing Group.
257. Kooth Digital Health (2023). Applied Ethical AI in Mental Healthtech.
258. Coote, L. *et al.* (2024). An early economic evaluation of Kooth, a web-based mental health platform for children and young people with emerging mental health needs. *Internet Interv.*, Vol 36, 100748.
259. National Institute for Health and Care Excellence (2022). 4 Economic evaluation | NICE health technology evaluations: the manual. *Guidance*. NICE.
260. Koutsouleris, N. *et al.* (2021). Multimodal Machine Learning Workflows for Prediction of Psychosis in Patients With Clinical High-Risk Syndromes and Recent-Onset Depression. *JAMA Psychiatry*, Vol 78, 195–209.
261. Jones, C. *et al.* (2023). Artificial intelligence and clinical decision support: clinicians' perspectives on trust, trustworthiness, and liability. *Med. Law Rev.*, Vol 31, 501–520.
262. O'Neill, O. (2002). A Question Of Trust.
263. Gould, M. (2020). Regulating AI in health and care. *NHS England Digital*.
264. Banerjee, S. *et al.* (2023). Involving patients in artificial intelligence research to build trustworthy systems. *AI Soc.*,
265. Banerjee, S. *et al.* (2022). Patient and public involvement to build trust in artificial intelligence: A framework, tools, and case studies. *Patterns*, Vol 3, Elsevier.
266. Greenway, F. T. *et al.* (2024). Hybrid mHealth care: Patient perspectives of blended treatments for psychosis. A systematic review. *Schizophr. Res.*, Vol 274, 1–10.
267. Kaur, D. *et al.* (2022). Trustworthy Artificial Intelligence: A Review. *ACM Comput Surv*, Vol 55, 39:1-39:38.
268. OECD (2024). Global Forum on Building Trust and Reinforcing Democracy.
269. O'Dell, B. *et al.* (2022). Building trust in artificial intelligence and new technologies in mental health. *BMJ Ment Health*, Vol 25, 45–46. Royal College of Psychiatrists.
270. Khera, R. *et al.* (2023). Automation Bias and Assistive AI: Risk of Harm From AI-Driven Clinical Decision Support. *JAMA*, Vol 330, 2255–2257.
271. Koutsouleris, N. *et al.* (2022). From promise to practice: towards the realisation of AI-informed mental health care. *Lancet Digit. Health*, Vol 4, e829–e840.
272. Toon, J. (2016). In Emergencies, Should You Trust a Robot? *News Centre*.
273. Trewin, S. *et al.* (2019). Considerations for AI fairness for people with disabilities. *AI Matters*, Vol 5, 40–63.
274. Bossewitch, J. *et al.* (2021). Digital Futures in Mind: Reflecting on Technological Experiments in Mental

Health & Crisis Support. University of Melbourne.
275. Saeidnia, H. R. *et al.* (2024). Ethical Considerations in Artificial Intelligence Interventions for Mental Health and Well-Being: Ensuring Responsible Implementation and Impact. *Soc. Sci.*, Vol 13, 381. Multidisciplinary Digital Publishing Institute.
276. Studman, A. (2023). Access denied? Socioeconomic inequalities in digital health services. Ada Lovelace Institute.
277. Corradi (2022). The role of technology in mental healthcare. The Nuffield Council on Bioethics.
278. Burr, C. *et al.* (2022). Trustworthy Assurance of Digital Mental Healthcare. Zenodo.
279. Zidaru, T. *et al.* (2021). Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. *Health Expect.*, Vol 24, 1072–1124.
280. Higgins, O. *et al.* (2023). Artificial intelligence (AI) and machine learning (ML) based decision support systems in mental health: An integrative review. *Int. J. Ment. Health Nurs.*, Vol 32, 966–978.
281. McCradden, M. *et al.* (2023). Evidence, ethics and the promise of artificial intelligence in psychiatry. *J. Med. Ethics*, Vol 49, 573–579. Institute of Medical Ethics.
282. Harrer, S. (2023). Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. *eBioMedicine*, Vol 90, Elsevier.
283. Coghlan, S. *et al.* (2023). To chat or bot to chat: Ethical issues with using chatbots in mental health. *Digit. Health*, Vol 9, 20552076231183542.
284. PARQAIR-MH (2023). PARQAIR-MH: A participatory initiative to include LGBTQIA+ voices in AI for mental health.
285. Kormilitzin, A. *et al.* (2023). A participatory initiative to include LGBT+ voices in AI for mental health. *Nat. Med.*, Vol 29, 10–11. Nature Publishing Group.
286. Hopkin, G. *et al.* (2024). Building robust, proportionate, and timely approaches to regulating and evaluating digital mental health technologies. *Lancet Digit. Health*,
287. Department of Health and Social Care (2024). Government response to the report of the equity in medical devices: independent review. *GOV.UK*.
288. Bragazzi, N. L. *et al.* (2023). The Impact of Generative Conversational Artificial Intelligence on the Lesbian, Gay, Bisexual, Transgender, and Queer Community: Scoping Review. *J. Med. Internet Res.*, Vol 25, e52091.
289. Skorburg, J. *et al.* (forthcoming). Ethical Issues in Text Mining for Mental Health. in *The Atlas of Language Analysis in Psychology*. (eds. Dehghani, M. et al.) Guilford Press.
290. Morgan, P. *et al.* (2024). Using artificial intelligence to address mental health inequalities: co-creating machine learning algorithms with key stakeholders and citizen engagement. *J. Public Ment. Health*, Vol ahead-of-print, Emerald Publishing Limited.
291. Buhmann, A. *et al.* (2023). Deep Learning Meets Deep Democracy: Deliberative Governance and Responsible Innovation in Artificial Intelligence. *Bus. Ethics Q.*, Vol 33, 146–179.
292. Involve Deliberative Public Engagement.
293. Reid, O. F. Public voices in AI Exploring how to represent public voices in AI research, development and policy.
294. Huang, K.-T. *et al.* (2024). The Influence of AI Literacy on User's Trust in AI in Practical Scenarios: A Digital Divide Pilot Study. *Proc. Assoc. Inf. Sci. Technol.*, Vol 61, 937–939.
295. Iserson, K. V. (2024). Informed consent for artificial intelligence in emergency medicine: A practical

guide. *Am. J. Emerg. Med.*, Vol 76, 225–230.
296. Vayena, E. *et al.* (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Med.*, Vol 15, e1002689. Public Library of Science.
297. Alexandre, F. *et al.* (2021). Why, What and How to Help Each Citizen to Understand Artificial Intelligence? *KI - Künstl. Intell.*, Vol 35, 191–199.
298. Elements of AI A free online introduction to artificial intelligence for non-experts.
299. Teachable Machine.
300. AI Myths Home.
301. DeepLearning.AI What machine learning can and cannot do - What is AI? *Coursera*.
302. Royal Society (2023). Science in the metaverse: policy implications of immersive technologies.
303. Nebeker, C. *et al.* (2019). Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Med.*, Vol 17, 137.

Contributors

POST is grateful to Hannah Gardiner for researching this briefing, to the Nuffield Foundation for funding their parliamentary fellowship, and to all contributors and reviewers. For further information on this subject, please contact the co-author, Natasha Mutebi.

Members of the POST board*

Aynsley Bernard, Kooth

Dr Graham Blackman, University of Oxford

Professor Adriane Chapman, The Governance in AI Research Group (GAIRG)*

Claudia Corradi, The Nuffield Council on Bioethics

Dr David Crepaz-Keay, the Mental Health Foundation*

Fiona Dawson, Mayden

Zoe Devereux, University of Birmingham*

Dr Piers Gooding, La Trobe University

Dr Caroline Green, University of Oxford

Lara Groves, Ada Lovelace Institute

Rachel Hastings-Caplan, Rethink Mental Illness and Mental Health UK*

Dr Gareth Hopkin, Science Policy and Research Programme Team, National Institute for Health and Care Excellence (NICE)*

Dr Becky Inkster, Cambridge University

Dr Grace Jacobs, Kings College London*

Lauren Jerome, Queen Mary University of London

Dr Caroline Jones, The Governance in AI Research Group (GAIRG)*

Dr Indra Joshi, Trustee for Lift Schools
Dr Andrey Kormilitzin, University of Oxford

Associate Professor Akshi Kumar, Goldsmiths, University of London

Professor Agata Lapedriza, Northeastern University; Universitat Oberta de Catalunya

Dr Paris Alexandros Lalousis, Kings College London*

Dr Sophia McCully, The Nuffield Council on Bioethics

Dr Rafael Mestre, Southampton University*

Associate Professor Stuart Middleton, Southampton University

Dr Max Rollwage, Limbic*

Dr Annika Marie Schoene, Northeastern University

Julia Smakman. Ada Lovelace Institute*

John Tench, Wysa

Dr James Thornton, The Governance in AI Research Group (GAIRG)*

Alli Smith, Office for Life Sciences

Mona Stylianou, Everyturn Mental Health*

Dr Pauline Whelan, CareLoop*

Dr Gwydion Williams, Wellcome Trust*

Dr James Woollard, Oxleas NHS Foundation Trust, NHS England*

Andy Wright, Everyturn Mental Health

Emeritus Professor Jeremy Wyatt, The Governance in AI Research Group (GAIRG)*

Information Commissioners Office*

The Joint Digital Policy Unit (a joint unit between the Transformation Directorate in NHS England, and DHSC)*

*denotes people and organisations who acted as external reviewers of the briefing. Some of them were also part of the interview contribution process. Note that contributors are listed in alphabetical order by surname.

The Parliamentary Office of Science and Technology (POST) is an office of both Houses of Parliament. It produces impartial briefings designed to make research evidence accessible to the UK Parliament. Stakeholders contribute to and review POSTnotes. POST is grateful to these contributors.

If you have any comments on our briefings please email post@parliament.uk. Please note that we are not always able to engage in discussions with members of the public who express opinions about the content of our research, although we will carefully consider and correct any factual errors.

If you have general questions about the work of the House of Commons email hcenquiries@parliament.uk or the House of Lords email hlinfo@parliament.uk.

DOI: https://doi.org/10.58248/PN737

Image Credit: Glenn Carstens-Peters on Unsplash

post@parliament.uk
parliament.uk/post
@POST_UK